

Durham Research Online

Deposited in DRO:

21 December 2016

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Heslin, Peter (2016) 'The dream of a universal variorum : digitizing the commentary tradition.', in Classical commentaries : explorations in a scholarly genre. Oxford: Oxford University Press, pp. 494-511.

Further information on publisher's website:

<https://global.oup.com/academic/product/classical-commentaries-9780199688982?cc=gblang=en>

Publisher's copyright statement:

This is a draft of a chapter that was accepted for publication by Oxford University Press in the book 'Classical Commentaries: Explorations in a Scholarly Genre' edited by Christina S. Kraus and Christopher Stray and published in 2016.

Additional information:

This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 640798).

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

The Dream of a Universal Variorum

Some years ago, the computer at my university library said that there were many Classics books on the shelves which were never being checked out and so were ripe for relegation to off-site storage. I went through the list of the on-shelf holdings with our liaison librarian and we sacrificed a few older monographs on unfashionable or obscure subjects in the hopes that this would appease the computer. The big problem was with the commentaries. Not the texts; for some reason, neither the computer nor the librarian had any quibble with these. Perhaps the stragglers and runts were protected by being part of a large and impressive herd like the Teubners or the Budé series. The value of having a text on the shelf to consult immediately was self-evident. For commentaries, the librarian wondered why we couldn't just call them up from off-site if we needed them. I explained that this would have been more than just a mild inconvenience: it would have made the commentary useless for many purposes. Often, we are not reading a text through, but are looking for explication for an aspect of a particular passage.¹ What we find in a commentary are parallels and other bits of evidence. So then we go and look at those parallel passages and at commentaries on those passages, and find further parallels which we look up and so on, jumping around until we find what we are looking for. If, at each stage in the process, we had to wait a few days for a commentary to come in from off-site storage, it would make the whole exercise unfeasible. I did, reluctantly, agree to relegate some older commentaries off-site, when there was a newer one on the same text. This, as we will see, is a mechanism of enforced artificial obsolescence which has become an essential condition of the form of the commentary.

Essentially, the corpus of classical commentaries constitutes a gigantic, decentralized, thickly cross-indexed reference work. But these do not look like reference books to a librarian, at least superficially, for they are a variegated jumble of volumes in different languages, from different publishers. Explaining all this to a sympathetic librarian made me realize what a remarkable and distinctive thing the corpus of commentaries on classical texts is. It has been constructed without any central authority, but it nevertheless manages to embed a countless number of cross-references. This magic is one of the greatest achievements of classical scholarship, though we do not often think about it. The web that holds together this sprawling reference work is constituted by the almost perfectly consistent set of methods we use to cite ancient texts. Citing the text of Plato by the pages of Stephanus' edition may not be the most convenient or logical thing to do, but we all tacitly accept it: the desirability of having one style of reference that everyone uses outweighs all other considerations. Even the most radical textual critic will pause before altering the standard system of citation for a text; many contemporary scholars approve of Lachmann's hypothesis that Propertius wrote five rather than four books of elegies, but everyone agrees that his edition's renumbering of the books was an unmitigated disaster, for it damaged the stability of the hard-won system whereby we can refer precisely and unambiguously to passages of classical texts.

Viewed in this way, the corpus of classical commentaries constitutes a hypertext *avant la lettre*. When one follows a parallel cited in one commentary to the text of another author, and then examines a commentary on that passage, and so on, the process is equivalent to clicking on hyperlinks from one webpage to another. The technology that underlies this hypertext is not http and html, but our total set of standard systems of citation for classical texts.² This is not a new observation; it lies at the heart of the seminal article on

¹On "the 'hit-and-run' commentary user," see Kraus 2002: 111.

²The equivalence is neatly illustrated by the Canonical Text Services project, which defines a digital protocol

the subject of the classical commentary and digital media published in 1999 by the late Don Fowler: “long before internet surfers clicked from link to link in an almost-endless chain of deferred pleasure, classical scholars were moving from Nisbet-Hubbard to Headlam to Mayor to Wakefield to Lambinus to Servius and back.”³ Fowler’s essay is full of optimism about two related possibilities: firstly, the creation of an enormous collaborative commentary “with infinitely large margins” on the whole of classical literature, written by the community of scholars, and secondly, the creation of websites would bring together references to all of the scholarly work on, e.g., a particular author.⁴ He did not know the name, but the kind of site Fowler was describing is now called a wiki. The year after his article appeared, a project called *Nupedia* was launched, which soon morphed into the collaboratively edited site called *Wikipedia*.⁵ In the light of how that project has developed, Fowler’s optimism about the glorious possibilities of collaborative on-line scholarship looks a bit naive.⁶ To be sure, *Wikipedia* is one of the greatest reference works ever created for facts which are amply documented and uncontroversial, if one uses it with due care. When it comes to interpretation, however, no one today would recommend its edit-wars, in which the palm of victory goes to the most persistent crank, as an exemplary model for scholarly communication.

In a sense, however, the second part of Fowler’s vision has come true, though not in the way he anticipated. He wanted a way to find and link to all of the current articles on a particular author or passage; this is now possible, more or less, thanks to projects like Google Books, Jstor, Persée and the like. Fowler’s article belongs to a pre-Google world in which it was still thought that human-edited indexes were the future of information retrieval; since then, Google has demonstrated comprehensively that full-text search trumps manual indexing (but see below for one major problem with this approach). So, a new lease of life has been granted to the traditional modes of academic production. The most important aspect of Fowler’s essay is not, however, its predictions, right and wrong, about the future; it is his insistence that digital opportunities demand a reconsideration of the form and content of the commentary. One conclusion Fowler draws in this regard is that the potentially infinite space available to digital publication ought to incline us toward presenting the reader with an overabundance of information rather than toward selectivity. In this essay, I intend revisit Fowler’s question of what infinite information capacity means for the form of the commentary, but from a different angle. My approach arises out of another development Fowler could not have foreseen, the advent of massive book-scanning projects, especially Google Books. When the librarian no longer needs to remove old commentaries from the shelf to make room for the new, can one justify the commentator’s infamously tralatitious habits?⁷ I am far less optimistic than Fowler was, and than Anderson in this volume is, about the claim that the immediate future of the commentary is digital. But this could be true for its past.

to resolve traditional textual citations: <http://cts3.sourceforge.net>.

³Fowler 1999: 441.

⁴Fowler 1999: 432, 441. For a more negative view of those margins, see Gumbrecht 1999: 452–3.

⁵Wikipedia contributors, “History of Wikipedia,” *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/w/index.php?title=History_of_Wikipedia&oldid=580554465.

⁶To be fair, Fowler admitted both the danger of trying to predict the future and the conservative effect of career structures on modes of publishing. And of course, there have been successful collaborative projects in Classics in the interim, such as the Suda On Line: <http://www.stoa.org/sol/>.

⁷On those habits, see Kraus 2002: 16–20 and *contra* Reeve 2004: 8.

DIGITAL FUTURE OR DIGITAL PAST?

In an ideal world, a digital commentary would make perfect sense. A reader would follow parallel passages not by pulling another text and another commentary off a physical shelf, but by simply clicking on a link to another digital text and another digital commentary. There are practical problems, however, which will make this a difficult ideal to realize in the near term. Most recent commentaries are protected by copyright and academic publishers are understandably reluctant to undermine their modest income streams by digitizing those works without ensuring a compensatory income stream via subscription-based access. The currently emerging scenario is that each publisher sells subscriptions for managed access to digitized commentaries along with their other classical content. The problem here is that this practice makes seamless linking from one publisher's commentary to another's very unlikely. Even hyperlinks between the same publisher's commentaries are unlikely, given the lack of an imperative to spend the necessary money. So we will probably end up with a system only slightly more convenient than the present printed page, in that we can consult most of our commentaries without heading to the library, but in most other respects, the hyperlinks will remain as cumbersome to follow, or even more so, than on the page. To follow up a parallel, you will need to go to the on-line library catalogue, find the e-book commentary, click through to the publisher's website and navigate manually the relevant page of the commentary. This is not a disaster, provided that you work in a well-funded ivory-tower institution that can pay for the subscriptions. If we leave the question of open access to the side, it is disappointing that, even for an amply resourced user, a process which ought to take a single click will continue to involve so much fuss. The separation of content into the websites of individual publishers means that there will never be a push for interoperability, unless we as a scholarly community insist upon it.

The alternative path toward a digital future for the commentary is via born-digital content distributed on-line as the principal mode of distribution. Here, the obvious problem lies in the structure of peer review, proofreading and quality assurance which has traditionally been mediated by publishers and which determines career paths for academics. Even if a commentary is written as digital content, distribution via a traditional publisher produces the same problems of interoperability. If an author publishes a digital commentary outside the patronage of a traditional publisher, who provides those services? There is an obvious role here for professional organizations, which already deal with those issues for many scholarly journals, and often do so in collaboration with traditional publishers.⁸ Such projects can serve an important role in setting standards for interoperability, but are unlikely to sweep away the traditional model of publishing commentaries. Perhaps, if professional organizations build the infrastructure to enable true hyperlinking of texts and commentaries, we can pressure publishers to adopt it also for their subscription-based content.

We may get to that digital nirvana eventually, but it does not seem imminent. Until that day, many of the benefits of the born-digital commentary will remain hypothetical. Much, and indeed, as I shall argue below, most of the value of a classical commentary lies in its role not as an arbiter of interpretation but as a collection of pointers to other resources. For a given passage, the commentator assembles data which can illuminate its meaning. In addition to the parallel passages with similar language or ideas which we have already discussed, these data may also consist of manuscript readings, works of art, material artifacts, historical or mythical persons or episodes and so on. Most of these are even harder to

⁸The American Philological Association is considering an initiative to organize the peer-review and publication of digital texts and commentaries for Latin, the Digital Latin Library Project: <http://apaclassics.org/publications-and-research/digital-latin-library-project>.

hyperlink than texts. Images of material artifacts bring even trickier questions of copyright, and there is a more fundamental problem in that our traditional systems of reference for artifacts are less well developed. Museums, libraries and archives may have some of their holdings on-line, but there is no standardized way of linking to, say, a particular item in a museum by accession number or to a particular page of a particular manuscript in a given library. There is no immediate prospect of any standardization in this regard, so a digital commentary which attempts to exploit its full potential at this stage is destined to become a collection of dead links within a few years.

The future of the commentary may be digital, but from the perspective of the present day, that future may be more distant and less shining than we would like to imagine. So let us turn our gaze in the other direction. If opportunities for digitizing new and recent commentaries are hamstrung by problems of copyright and the role of traditional publishers, none of these apply in the case of older material which is out of copyright. As classicists, we are sitting upon a vast wealth of commentary material which is being digitized in a completely haphazard and unsuitable manner. Anyone reading a digital version of a classical text ought to be able to view at a single click the notes of every commentator from the incunabular age to the watershed of copyright. But we have not evolved structures which will reward those who work toward such an outcome. Before we sort out our future, we might try to sort out our past.

The material which is out of copyright is not to be sneezed at. As far as student commentaries are concerned, Victorian texts aimed at secondary-school students who had been studying Latin since primary school are often quite suitable for contemporary university students who have been studying the language for fewer years. With respect to scholarly commentaries, the out-of-copyright material is even more valuable. The comparative material brought to bear by even the most up-to-date commentary usually exists already somewhere in the tradition. Right from the very beginning of the classical commentary, when books were expensive and rare, the Dutch did a great service in putting together the original variorum commentaries. Making available to a new audience the insights of an earlier generation of scholars is not a new or eccentric idea. Indeed, it is such an obvious desideratum that it is most strange that there has been little effort in this direction so far. There is nothing new about librarians removing old books from the shelf to make room for the new. The novelty is that in the digital world, there is no limit on the space available on the shelf. There should no longer be a necessity for the new commentary to reiterate what the old one said. We can simply direct the reader to the original note. But we do not do this. We continue to write commentaries as though books were rare and information hard to come by. We have not digitized the historic commentary tradition, and so we can continue to pretend that the job of the commentator is to act as the sole and definitive arbiter of what the reader needs to know. But technology has moved on.

THE UNIVERSAL VARIORUM

There are three different models, at increasing levels of funding and, therefore, improbability, for imagining a digitization project for the corpus of past commentaries. The most rudimentary is an indexed collection of links to scanned books which are nothing more than digitized page images. The user would have to flip through the pages manually. At a higher level of ambition, these page images would have added to them good enough uncorrected Latin/Greek optical character recognition (OCR) to permit viable searching. A further level of technological sophistication would include automatic recognition of lemmata in such a way as to permit the reader to jump from the digital text directly to the

corresponding lemma in each commentary upon that passage. These three levels of implementation require three different kinds of technological intervention.

There is a technological hurdle which needs to be overcome even for the simplest form of harvesting digitized commentaries: it has to do with metadata. At the moment there are a number of large-scale digitization projects of printed books under way, such as Google Books and Gallica. In theory, we ought to be able to wait for these large, general projects to scoop classical commentaries into their maw. One problem is that some of these projects seem to place older books and books in Latin and Greek on a low priority. Furthermore, the poor quality of metadata for these scans can make them very difficult to find. Google's motto seems to be "Who needs metadata?", much to the consternation of librarians who have traditionally spent so much time and effort collecting it.⁹ It is true that Google has demonstrated that, in most domains, full-text search is vastly more powerful and convenient than elaborate manual cataloguing. But this policy has been disastrous for some kinds of works. For example, it is presently near to impossible to find a particular article in a run of old journals digitized by Google if all one has is a citation. If you happen to have the title of the article, you might get lucky with full text search, but even that is quite fragile.

The failure of Google to harvest full metadata when digitizing library collections is equally problematic for commentaries. Google may only know that the author of a book is Horace, the title is *Opera*, and the year of publication; it often does not know if it includes a commentary or the identity of the editor and/or commentator. For a well-known author, especially one popular as a school text, there may be hundreds of different books with almost identical metadata. When these files are uploaded to a site that aggregates free media, such as the Internet Archive (archive.org), they join hundreds of others, which may have more metadata, but without any consistency. Where the modern editor or commentator is given, it is sometimes listed as a second author, sometimes as part of the title, sometimes elsewhere.

A minimally useful corpus of digitized classical commentaries would consist of an index of ancient authors and works, with links to scanned texts of scholarly or student commentaries at sites like the Internet Archive. Because of the problems with metadata, an index like this would have to be compiled by hand, but it would be genuinely useful. Its usefulness will increase as more and more old books are scanned. A few years ago, there were few early commentaries to be found on-line; for example, Bentley's Horace was available on the Internet Archive, but only because it had been scanned from a modern reprint. Today, however, many, though not all, commentaries previously confined to the early printed book room can be downloaded from Google Books and other libraries with major digitization initiatives, such as those in Paris, Heidelberg and Munich. For example, a quick survey of the commentaries on Propertius listed in Thompson's entry for that author in volume 9 of the *Catalogus Translationum et Commentariorum* indicates that one can download from Google Books the most important commentaries from the sixteenth century and after (Scaliger, Passerat), while the fifteenth-century commentaries (Calderini, Beroaldo) are still out of reach, except insofar as they are excerpted in later variorum editions. Still, this is progress, for a few years ago I was not able to find scanned copies of either Scaliger or Passerat. So we may hope that such digitization projects will continue to proceed backward in time though the world of incunabula.

The limitations of a simple index of links to files with scanned page images are considerable. Its books are not searchable and must be paged through to find the required content. Flip-

⁹HathiTrust is running a project designed to reunify the digitized versions of library books with their metadata: <http://www.hathitrust.org>.

ping through a PDF, even on a fast computer, is not a very efficient way to locate content. Physical books will probably never be bettered in that regard. For commentaries, this kind of usability matters. Since one usually wants to read a short note on a particular passage, if the effort to find that note is too great, the reader may simply move on to a more accessible commentary. The first step necessary in making the corpus of historic commentaries accessible is to improve the quality of OCR for Latin and ancient Greek, in order to make it possible to search for particular phrases. Google has demonstrated that it is not necessary to have 100% accuracy for search to be useful. Text recognition which is “dirty”, or uncorrected by human hands, is sufficient at 95% accuracy for rough-and-ready searching. The key point is that the reader looks at the page image, so the extracted, searchable text does not have to be a perfectly correct substitute. This is crucial, for uncorrected OCR can be done automatically, whereas correcting by hand is slow and expensive.

There are a number of issues that need to be addressed when performing OCR on Latin or Greek texts. There must be a word-list of legitimate forms, and a means of dealing with ligatures and old-fashioned typographical practices. There must also be a method for expanding the abbreviations which are so common in early printing. These are very far from being intractable problems, and efforts are under way to solve them.¹⁰ Fortunately, almost all Latin books were printed in italic or roman typefaces, so we can ignore the problem of blackletter printing. For Greek, the biggest problem is in the vast number of ligatures employed in older typefaces of the *Greco du roi* type. This is not a problem different in kind from what is routinely encountered in OCR technology for modern languages, such as for Arabic ligatures. Implementing good enough OCR for old books in Latin and Greek is not a matter of creating new technological solutions; it simply involves applying well-understood technologies to the conventions of those books.¹¹

If one were to develop good enough open-source OCR for Latin and Greek, one would hope that large-scale scanning projects would begin to use it. There is something tragic about the way projects like Google Books and Gallica often embed text in their scanned books which is simply gibberish. Presumably, they would not object to doing it better, and if classicists presented them with a tool that worked, they would surely start to use it. Even if not, their efforts in scanning the books would not be in vain. We could take their PDFs, remove the gibberish, re-do the OCR, embed the improved text, and upload the improved PDF to the Internet Archive. This process could be automated.

This is a matter which has implications well beyond the classical commentary, and indeed well beyond classics as a discipline. The lack of robust, affordable and usable open-source OCR for Latin is having a tragic effect on the digitization of a sizeable part of our intellectual patrimony. From the Renaissance until the end of the eighteenth century, the vast majority of books printed throughout Europe were written in Latin. It was the *lingua franca* of every intellectual discourse: the natural sciences, mathematics, philosophy, theology, law, literary criticism, geography, archaeology, music, medicine. Will the vast heritage of early printed books in Latin survive the transition to digital? As libraries shrink their shelf-space and push books farther off-site, we will need to rely more and more on digital surrogates. Yet, at the moment, the surrogates we are creating cannot be searched through and have minimal metadata attached to them. That is a disastrous combination for scholars in many fields beyond Classics. If we classicists develop tools to properly digitize our own Latin heritage, we will be doing a great favor to the study of the Western tradition as a whole.¹²

¹⁰See Rydberg-Cox 2009.

¹¹See White 2012.

¹²A new, major effort to address these fundamental issues is the Open Philology project: <http://sites.tufts.edu/>

A collection of searchable scanned texts would be better than mere page images, for it would give you a way to try to find the passage of interest to you without flipping through the whole file. But it still falls short of the ideal. What we really want to be able to do is to call up with one click the correct page of the scanned book for a given citation of an ancient text. With some thought, one ought to be able to write a program which would, given a classical text and a scanned commentary, be able to identify most of the lemmata in the latter. It would have to take into account that the given text and the text of the commentary may differ or appear out of order, and that the OCR will not be perfect. Nevertheless, it is a problem should be solvable in a high enough percentage of cases to produce an index to the PDF which would permit you to jump, most of the time, from the citation to the correct page of the commentary. In this way, we could transform the quasi-hyperlinks implicit in the standard systems of citation employed in the corpus of classical commentaries into a true hypertext.

Consider a digital environment for reading classical texts where every phrase could bring up a menu of historic commentaries, and selecting one brought you right to the relevant page for the passage you are reading. This would truly be a universal variorum commentary (at least up to the watershed of copyright), and such an environment would not be so difficult to construct. The usual reasons given for failure to advance a digital humanities project of this kind are: lack of funding, lack of organization and lack of understanding of the possibilities on the part of traditional humanists. All of these are probably to blame, but I wonder if by blaming externalities we are letting ourselves off the hook. Is it possible that it suits the contemporary writer of the classical commentary for historic commentaries to remain difficult of access?

The notoriously accumulative nature of the commentary is rooted, as a historical phenomenon, in the scarcity of books and of shelf-space. The early variorum commentaries were welcome, despite being acts of blatant appropriation, because books were so difficult to access in that age. Commentaries ever since have tended to aggregate and condense older material, in the knowledge that the new commentary will soon push the old commentary off the shelf and into off-site storage, as part of the process of thinning that was mentioned at the start of this piece. Today, our digital shelves are infinitely capacious, but we continue to write commentaries as though the situation has not changed, as though the commentator will be the reader's only resource. To my mind, the telling failure of the commentary in the digital age is not the missing of opportunities for shiny new content, but the failure to realize the possibility of a new relationship between the reader of a commentary and the commentary tradition as a whole and hence a new rhetoric of commentary-writing.

TEXTUAL CRITICISM CONSIDERED AS A MENTAL DISORDER

Essays on the digital commentary have tended to ignore the question of the accompanying text, but Fowler's injunction to examine the form of the commentary in the light of digital possibilities is just as important for classical texts. The writing of commentaries cannot be separated from the editing of a text: the one implies the other, even if the commentary is printed without a text, or with one adapted from another edition. The present moment ought to promise a new dawn for textual criticism. It is finally possible to offer the reader not a single text with variants confined to an apparatus, but a text that changes in response to the reader's experimentation. On this view, the job of the editor should now be not to decide on the text, but to marshal all of the evidence in such a way for the reader to manipulate conveniently. The reader ought to be able to see instantly the text as reported

perseusupdates/2013/04/04/the-open-philology-project-and-humboldt-chair-of-digital-humanities-at-leipzig/.

by any given witness or previous editor, not as a collection of variants reported against the editor's text, but in its own right. The editor could still give his or her preferred text, but as one option among many, which the reader could change at will.

The usefulness of the current generation of digital classical texts is severely limited by the general absence of critical apparatus. Whether this omission has been due to technological barriers or fears about the copyright status of a modern editor's apparatus, the result has tended to produce a generation of digital readers who are no longer used to confronting at every turn the evidence of our uncertainty about the transmission of those texts. There have been a number of individual efforts to remedy this situation: there are, for example, very full on-line textual resources for Catullus and Horace.¹³ A much broader but necessarily less full effort to remedy the absence of an apparatus has been made by the *Musisque Deoque* project for its collection of Latin poetry.¹⁴ There are serious questions about the interoperability and sustainability of these isolated efforts, and they all operate within the old paradigm of an editor's text with variants relegated to a human-readable apparatus, but they are a step in the right direction. Such projects tend to apply Fowler's principle of "infinitely large margins" to the apparatus, removing the need for selectivity; ideally, readings would link back to images of the underlying MSS.¹⁵ Totally comprehensive digital editions like these may be ideal when genuinely conceived *editorum in usum*, but may overwhelm the more casual reader. The ideal situation would be to permit the reader to decide how much detail she would like to see. The reader would then tell the computer what to show: the text of a particular witness, of a particular editor, or the apparatus of a particular editor. This would require, however, producing a marked-up digital apparatus which is designed to be read not by people but by machines.

Modern digital editions of classical texts tend to take for granted the old notion that editor's central task is to establish a single, stable text. One exception appears to be the Homer Multitext project, which embraces the potential of digital technology to destabilize the idea of a canonical text.¹⁶ The difference for Homer is, of course, that there is a large body of scholars who believe that the labile nature of the Homeric poems is an essential feature of their nature rather than a result of corruption. The existence of the Homer Multitext project renders the decision to produce a traditional edition of Homer a polemical act, which is as it should be. But how different is Homer from literary texts? Criticisms of the stemmatic method usually focus on the practical difficulties: that in practice most traditions are open or contaminated or both, which limits the extent to which one can eliminate witnesses from consideration.¹⁷ But there is a much larger issue, which textual criticism will have to take more seriously in the light of the possibilities of a digital age than it has been accustomed to doing in the age of print. Does the hoary but convenient positivist fiction of a stable and recoverable authorial archetype still make sense for any ancient text? Textual critics working on modern texts with much simpler dynamics of transmission have already abandoned this ideal as romantic and ahistorical.¹⁸

The practical considerations of hot-metal typography mean that printed editions of ancient texts have been forced to take the form of a single text, with variants recorded separately.

¹³See Catullus Online (www.catullusonline.org) and Repertory of Conjectures on Horace (<http://www.rekstlab.uio.no/horace>).

¹⁴<http://www.mqdq.it/>.

¹⁵This is not a new idea; Fowler 1999: 437 points to the digital *Canterbury Tales Project*. The Catullus Online project mentioned just above provides links to manuscript images.

¹⁶<http://www.homermultitext.org/>.

¹⁷See Reynolds and Wilson 1991: 214–6.

¹⁸See McGann 1983.

The modern editor of a printed edition implicitly presumes that the ancient author experienced publication in much the same way: via a manuscript which took a definitive form when it was published. We know, of course, that the informal circulation and performance of literature in antiquity was as important as distribution through booksellers. Anyone who has put a document on the web knows the irresistible temptation to fiddle with it post-publication. We have swiftly moved into a new era which makes it apparent that the age of the printing press was an aberration – a very long aberration – from the norm in which there was no fixed moment of publication. When we write digital texts, we are like ancient authors, in that we have no reason to stop making small changes whenever we come back to them. It is an article of faith among textual critics that variants can in general be classified as errors or true readings. In the ancient world, where the author, like the creator of a web-page, had no reason to stop making changes to his text, it would be surprising if those authorial variants were not transmitted so as to appear in the manuscripts we possess. Why do editors rush to attribute plausible variants to meddling third-party interpolators?¹⁹

The presence of ubiquitous, genuine authorial variants in our witnesses would not be a congenial scenario for a traditional critic, for whom the determination of agreement in error is the primary tool for constructing the stemma. If variant readings are allowed to go back to the author, the situation gets much messier. This is a more serious problem than when, as often happens, the editor comes to the conclusion that the tradition is open or contaminated. The collapse of the stemma into an unstructured web of relationships does not mean, however, that one cannot construct a graph which maps out the relationships among witnesses. The so-called “New Stemmatics” (by now several decades old) is a purely digital approach to the problem, which makes use of computer-based analysis of as many witnesses as possible to determine their web of relationships.²⁰ This new approach is in part a reaction to Bédier’s devastating demonstration (at least for medieval texts) that the allegedly scientific stemmata constructed by editors over the centuries have been determined more by the subconscious prejudices and implicit presuppositions of those editors than by the nature of the evidence.²¹

At the core of this new methodology is the decision to classify witnesses by means of all variants, not by tendentiously-selected instances of purported agreement in error. This is necessary because a computer cannot tell the difference between an error and a variant. But at least the computer, unlike the textual critic, is honest about the limitations of its intelligence. The result of an algorithmic analysis of textual variants will therefore be a web of relationships rather than a stemma which hangs vertically from an alleged archetype. Computational techniques can group manuscripts together in families without any prior prejudice as to which reading is closest to the archetype. This methodology is clearly superior to the inherently circular proceeding of the traditional critic, who decides that one variant is probably what the author wrote while another is an error, and uses that decision to support the priority of one witness over the other; that priority is used elsewhere to support its readings, and so on.²² Some variants are, of course, obvious scribal errors that could not possibly go back to the author. But many others, despite the black-and-white rhetoric of obvious truth and abject error which has become the textual critic’s inevitable

¹⁹See the overview of Tarrant 1989.

²⁰See <http://www.textualscholarship.org/newstemmatics/index.html>.

²¹On the reaction by classical textual critics to this demonstration, see Timpanaro 2005: 157–87. For a skeptical response, see Reeve 1986 with Timpanaro 2005: 207–15.

²²On the logical fallacy entailed in making judgements of truth and error prior to determining the relationship of the manuscripts, see Robinson and O’Hara 1993. Compare the dismissive remarks of Reeve 1986: 58 with n. 2.

mode of discourse, are not. There are thus two problems with traditional practice. The first is methodological: determination of the relationships between witnesses should come prior to any judgement about what witnesses are closest to what the author wrote, not afterward. A stemmatic method based upon first determining “agreement in error” is further flawed by this insistence on premature separation of sheep from goats. When your method requires you to begin from a judgement about error and truth, your text is unlikely find much room for the possibility of genuine variants surviving from the author’s hand.

The usual explanation for the presence of plausible variants in our witnesses is in the working of copyists: omission, interpolation, introducing glosses, and attempting to make good losses to the text through conjecture. There is evidence for all of this activity on a small scale, but the general unwillingness of editors to accept that very many of our variants may go back to the author seems perverse. Anyone who has thought for more than a minute about the nature of publication in antiquity must see that there is every reason to expect that the author continued to revise his text over the course of its initial period of distribution. The obvious exception is where the author was dead and publication was posthumous. Take the *Aeneid*, for example. This is the closest thing we can find in antiquity to publication in the definitive and final sense it acquired in the age of printing. Yet even here, there were non-canonical variants circulating in antiquity which could pass, in the view of intelligent contemporary readers, as passages from Virgil’s autograph MSS which were excised by the editor(s) before the posthumous publication of the epic. The efforts of critics to prove on stylistic grounds the interpolation in late antiquity of passages like the Helen episode tend rather to demonstrate the hopeless subjectivity of this methodology and the narrowness of the available comparative linguistic evidence.²³ Ovid and Lucan thought these were interesting potential variants, and that should be enough to make them interesting variants for us.²⁴ But the positivist bias of textual scholarship has obscured an important uncertainty principle: the activity of a sufficiently talented and early interpolator is from our perspective indistinguishable from authorial revision.

If even the *Aeneid*, despite the death of its author and the alleged *imprimatur* of Augustus upon the canonical text, could provide a posthumous stream of plausible variants, it is vastly more likely to have happened to less stable texts with more fluid and informal histories of dissemination. A contemporary of Virgil would often have recited his poetry for friends and patrons privately, would have given public recitations, and would have urged his patrons to send copies of his work to other influential men of letters, who might in turn make copies for others and so on. The work might eventually be deposited in a library or copied by a booksellers, who generally did not compensate the author. Over the course of this extended period of performance and publicity for the new work, the author would naturally have been fiddling with it. At a certain point in that process, once the work was polished enough to be deemed ready for public circulation, the author would have authorized general copying, either implicitly or explicitly.²⁵ But he had no financial or other interest in freezing that version absolutely. Knowing that his text was immediately subject to errors of reproduction, he could have continued to make his own small changes without destabilizing it any more than was already inevitably the case. If he were, for example, to give a public recitation of a new work after authorizing its initial public copying, it is hard

²³For a hypersceptical discussion, see Goold 1970; for a more appreciative view of the talent of the alleged interpolator, see Horsfall 2006–7.

²⁴The alternative opening lines of Virgil’s epic were known by Ovid: see Conte 1986: 84–7. The Helen episode was known by Lucan: see Bruère 1964; *contra*, Murgia 2003 argues that the alleged interpolator was influenced by Lucan.

²⁵See Starr 1987.

to imagine that he would have felt any qualms about continuing to modify it in places.

In their discussion of the “limitations of the stemmatic method”, Reynolds and Wilson do go so far as to consider cases where there might have been authorial corrections, but their language is telling: “Sometimes these would be extensive enough to justify us in speaking of a second edition.”²⁶ Here we see the dominion of the alien conceptual framework for publication imposed by the modern experience of hot-metal typography. There were indeed a few true “second editions” of some ancient works, but that is a separate issue. Most of our texts surely went through a period when authorized public copying and continuing authorial revision overlapped. This need not have prompted the author to issue a new, discrete “edition” except in circumstances where the changes were unusually extensive or important.

The belief in a stable, recoverable authorial archetype, or at least in an authoritative edition produced in late antiquity or perhaps even the Carolingian period, has never needed to be defended very much, for the simple reason that the printed form of the text demanded it, forcing the editor to make a choice of what to put in the text and what to relegate to the apparatus. There was little point in arguing that our texts were already a horizontally contaminated mess in their authors’ lifetimes in the days before we had a way of representing for the reader such a fluid and provisional state of affairs. The availability of more dynamic and interactive digital possibilities for representing the text has the potential to clarify the proper role of the editor, who presents the evidence for the text to the reader, as opposed to the textual critic, who problematizes our assumptions about that text. The value of the editor lies in the apparatus, not the text. The gathering and organization of evidence and weighing of its importance is vastly more important than the decision about what text to reproduce by default, which is a matter of opinion on which all readers will routinely disagree and which in digital media ought to be a parameter that the reader can change. In other words, the editor should not be presenting a text but many texts, a multitext. But the traditional form of the printed book forces the editor to put such an importance on the single, established text that this choice has often seemed to be his or her main business, and so we have conflated the separate roles of the editor and the textual critic. The textual critic can and probably should be provocative and opinionated; the editor must be scrupulously neutral, transparent and non-judgmental. This confusion of roles has fatally poisoned the rhetoric of philology. It must be the only discipline of the liberal arts in which the scholar is expected to speak in terms of absolute truth and absolute falsehood rather than of probability, ambiguity and provisional, relative truth.²⁷

The specious imperative of choosing one and only one text to print transforms textual criticism from a set of problems over which reasonable people might disagree into a test of the personal authority of the critic; disagreement rapidly descends into puerile, *ad hominem* abuse. The ill-founded certainty of textual critics in matters better treated as deeply ambiguous and finally undecidable alienates them from those readers who would rather be trusted to make up their own minds about the problems presented by the text. An editor who removes a suspected interpolation may seem to do unforgivable violence to the author; if instead the reader were presented with a clickable button which permitted him or her to experiment with seeing the text with or without the suspect words, how much less a potential act of vandalism might this seem? How much less nastiness might there be in our little world? In its more toxic manifestations, textual criticism looks not so much an intellectual discipline as a mental derangement. Much of this is due to the inflexibility

²⁶Reynolds and Wilson 1991: 215.

²⁷On the corrosive effects of that style of discourse on humanism, see Sullivan 1962.

of moveable type and the tyranny of the printed word, a situation alien to antiquity, and, increasingly, to us.

CONCLUSION

The potential of the commentary as described by Fowler to embrace a more open and less positivistic epistemology in the digital world seems as distant as ever. The traditional forms of commentary-publishing continue and editors continue to pursue the chimera of a single, stable authorial autograph. In the past ten years, the biggest change in the landscape of digital texts and commentaries has been the increasing availability of scanned versions of out-of-copyright editions. Inadequate though these scans are, they mark a major structural change; as Fowler insisted, such changes should lead us to question whether the forms of our production need changing. When shelf-space is no longer a consideration, why does a new commentary feel compelled to repeat the evidence presented in the old ones? When books were scarce, the early Dutch publishers did the world a great service with their variorum editions in bringing hard-to-access commentary material to a new and wider public: they were the Google Books of the seventeenth century. At least they came by their borrowings honestly. The subsequent history of the commentary has too often consisted of notes which essentially repeat the cumulative wisdom of preceding ages, with perhaps a small addition. When the commentator disagrees with that previous tradition, the constraints of space in the printed book have tended to produce a one-sided note in which the other interpretation is given short shrift. A less ambitious, but more achievable, goal than Fowler's for the commentary in the digital age would be one in which the tralatitious material is finally ejected. We may dispense with the fiction that there is only room on the shelf for a single commentary, with the librarian regularly ejecting the obsolete into off-site storage. A meta-commentary could hyperlink to earlier discussions without repeating their content, and, where appropriate, comment upon those comments. It ought to be more possible now to realize Fraenkel's vision of the commentator as starting from the total history of interpretation of a passage.²⁸ Or, if that is too sweeping an ambition, at least we may keep the commentator honest about his or her borrowings.

If we were to build a universal multitext and variorum, encompassing the full history of the constitution of the text by all of its editors, along with all of its witnesses, and including a full history of its commentary, we would be handing control of the reading process to the reader. Whether there is sufficient reward for the scholar in building such a tool remains to be seen. At the very least, it should not be too much to ask that we begin to organize and improve the raw scans of old classical editions and commentaries that Google and others are currently eructing in a disgraceful mess. Once that is done, if they are linked together and hyperlinked to an electronic text, the reader will have at his or her disposal a fairly universal variorum edition of all texts and commentaries before the watershed of copyright. If publishers can eventually be induced to play along, we might even add subscription-based hyperlink access to in-copyright commentaries. If future commentators know that all readers already have that mass of information at his or her fingertips, it might focus their attention on areas where there is genuinely new evidence to adduce.

BIBLIOGRAPHY

Bruère, R. T., 1964. "The Helen Episode in *Aeneid* 2 and Lucan," *Classical Philology* 59, 267–8.

²⁸There are other difficulties with this ideal, as demonstrated by Goldhill 1999:409–10.

- Conte, G. B., 1986. *The Rhetoric of Imitation: Genre and Poetic Memory in Virgil and Other Latin Poets*, ed. C. Segal (Ithaca: Cornell University Press).
- Fowler, D. P., 1999. "Criticism as Commentary and Commentary as Criticism in the Age of Electronic Media" in *Commentaries - Kommentare*, ed. G. W. Most (Göttingen: Vandenhoeck and Ruprecht), 426–42.
- Goldhill, S., 1999. "Wipe Your Glosses" in *Commentaries - Kommentare*, ed. G. W. Most (Göttingen: Vandenhoeck and Ruprecht), 380–425.
- Goold, G. P., 1970. "Servius and the Helen Episode," *Harvard Studies in Classical Philology* 74, 101–68.
- Gumbrecht, H. U., 1999. "Fill Up Your Margins! About Commentary and Copia" in *Commentaries - Kommentare*, ed. G. W. Most (Göttingen: Vandenhoeck and Ruprecht), 443–53.
- Horsfall, N., 2006–7. "Fraud as Scholarship: The Helen Episode and the Appendix Vergiliana," *Illinois Classical Studies* 31/32, 1–27.
- Kraus, C. S., 2002. "Introduction: Reading Commentaries / Commentaries as Reading" in *The Classical Commentary: History, Practices, Theory*, ed. R. K. Gibson and C. S. Kraus (Leiden: Brill), 1–27.
- McGann, J. J., 1983. *A Critique of Modern Textual Criticism* (Chicago: University of Chicago Press).
- Murgia, C. E., 2003. "The Date of the Helen Episode," *Harvard Studies in Classical Philology*, 101, 405–26.
- Reeve, M. D., 1986. "Stemmatic Method: 'qualcosa che non funziona?'" in *The Role of the Book in Medieval Culture*, ed. P. Ganz (Turnhout: Brepols), vol. 1, 57–70.
- Reeve, M. D., 2004. "Snow on Cithaeron," *Classical Review* 54, 5–12.
- Reynolds, L. D. and Wilson, N. G., 1991. *Scribes and Scholars: a Guide to the Transmission of Greek and Latin Literature* (Oxford: Clarendon, third edition).
- Robinson, P. and O'Hara, R. J., 1993. "Report on the Textual Criticism Challenge 1991," *Bryn Mawr Classical Review* 03.03.29 (<http://bmcr.brynmawr.edu/1992/03.03.29.html>).
- Rydberg-Cox, J. A., 2009. "Digitizing Latin Incunabula: Challenges, Methods, and Possibilities," *Digital Humanities Quarterly* 3.1 (<http://www.digitalhumanities.org/dhq/vol/003/1/000027.html>).
- Starr, R. J., 1987. "The Circulation of Literary Texts in the Roman World," *Classical Quarterly* 37, 213–23.
- Sullivan, J. P., 1962. "The Leading Classic of His Generation," *Arion* 1, 105–22.
- Tarrant, R. J., 1989. "The Reader as Author: Collaborative Interpolation in Latin Poetry" in *Editing Greek and Latin Texts*, ed. J. N. Grant (New York: AMS), 121–62.
- Timpanaro, S., 2005. *The Genesis of Lachmann's Method*, ed. and trans. G. W. Most (Chicago: University of Chicago Press).
- White, N., 2012. "Training Tesseract for Ancient Greek OCR," *Tò Εὐτυπον* 28–9, 1–11 (<http://www.eutypon.gr/eutypon/pdf/e2012-29/e29-a01.pdf>).